Hand Vasculature Image Analysis and Gender/Age Classifications

Amr M. Ashour, Mona A. Hussein
Ahmed M. Badawi
Systems and Biomedical Engineering Department, Cairo University, Giza, Egypt

***Abstract*— Objective: Biometrics is the science of identifying individuals based on physiological and behavioral characteristics. In this paper, we used dorsal hand veins pattern for age and gender classification. Materials and Methods: First, we used a dataset of 1000 images for 160 persons (for age and gender) that were acquired by a charge-couple device (CCD) monochrome camera. Then, we applied morphology, bilateral and median filter to reduce the noise, and contrast limited adaptive histogram equalization (CLAHE). We applied local binary pattern (LBP), local ternary pattern (LTP), and gray level cooccurrence matrix (GLCM) with its properties (dissimilarity, correlation, contrast and ASM). After that, we used k nearest neighbor (KNN) and support vector machine (SVM) classifiers. We create a fusion between classifiers and features. Results: We got best results from the fusion of KNN and SVM classifiers in age and gender classification. In gender, the best result was 85% in 50% data set with best K=35 with LBP and the best result in 70% data set was 85% with best K=13 with ASM. In age, the best result in 50% data set was 85% with best K=41 & 47 with LBP and LTP respectively and the best result in 70% data set was 86% with best K=41 & 47 with LBP and LTP respectively. We also got best results from the fusion of GLCM features in age and gender classification. In gender, the best result in KNN with GLCM was 82% and K=3. In Age, the best result in KNN with GLCM was 87% with best k=21. Conclusion: In this paper, we presented a biometric technique to classify age and genders of individuals using dorsal hand veins. Significance: create an efficient system over current security practices.**

Keywords— Dorsal hand vein, biometrics, age and gender classification, morphology filter, bilateral filter, median filter, contrast limited adaptive histogram equalization

I. INTRODUCTION

Biometrics is related to identifying individuals based on physiological and behavioral characteristics. Physiological could be facial image, voice pattern, eye recognition (Iris), and finger print. There are several security systems based on such characteristics, but, unfortunately, they are not strong enough. Qinghan [1] said that recent research has revealed that biometric technologies can be defeated. Since biometrics are not secrets, there exists a risk of them being captured, copied, and forged.

Recently, dorsal hand veins pattern has proved that it contains unique information as we can use them not only for authentication but also for identifying the range of age and gender. It is a great biometric system as it is fake-resistant, unique, user friendly, and hard to be damaged.

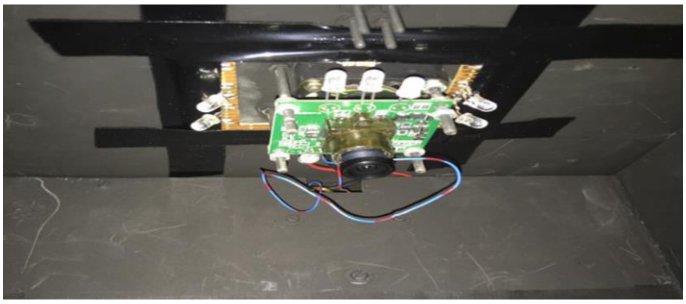

Fig.1 the camera with surrounding LEDs

Vein pattern uniqueness and low-cost authentication system was proposed by Shahin et al [2] as they use fast correlation of near infrared. Amayeh et al [3], proposed gender classification based on hand shape by the geometry of each hand. Age and gender classification based on finger vein pattern by using local binary pattern (LBP) was proposed by Damak et al [4]. In [5], McFadden and Shubel build a system able to recognize the age range and the gender of individuals from their venous network characteristic. Wang et al [6] proposed a new identification method of finger vascular patterns using a weighted LBP and support vector machine (SVM).

Wang et al [6] analyzed hand dorsal veins and classify age based on differences in blood flow and skin state by using gray scale histogram and mean pixel gray value of the vein and skin as features. They used K-means clustering classifier. Calin et al [7] analyzed human dorsal hand skin texture using hyperspectral imaging technique for assessing the skin aging process, and he mainly used gray level cooccurrence matrix(GLCM). Kröger et al [8] has proved that gender, body mass index, age and varicose veins influence cross-sectional area.

In this paper, we explain how we can use such pattern for age and gender classification. We will discuss the first part of materials and methods, which is data preparation.

II. MATERIALS AND METHODS

A. *Data Preparation*

We used a dataset of 1000 images that represent the left hand of 160 persons (80 males and 80 females, 110 young and 50 old) that were acquired by a charge-couple device (CCD) monochrome camera (see fig.1) that provide thermal image and relatively low cost to thermal camera as it is highly sensitive for near infrared (IR) spectrum covering 700-

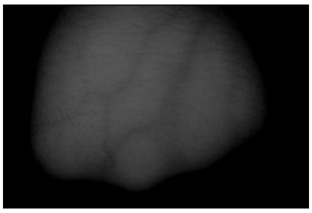

Fig.2 Original Image

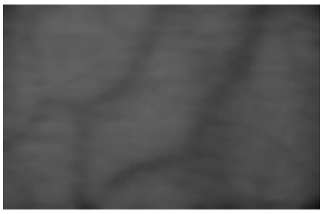

Fig.3 After CLAHE

1400nm and it was surrounded by 24-IR LEDs mounted in square shape 6-in each side. Each person had from 5 to 7 images and few of them had 10 images. In age classification, we consider the young people as below 35 years old and the old people as above it.

We made two experiments. In the first one we used 70% of the data as a training set as it represents 112 persons or 700 images. This represents 56 males (350 images) and 56 females (350 images) in gender. This also represents 77 young (560 images) and 35 old (140 images) in age. In the second one we used 50% of the data as a training set as it represents 80 persons (500 images). This represents 40 males (250 images) and 40 females (250 images) in gender. This also represents 55 young (400 images) and 25 old (100 images) in age.

B. Image Enhancement

First, we applied Otsu thresholding, as it aims to automatically find optimal threshold for image binarization to find best Region of Interest (ROI), which means applying the best segmentation for the original image. Then, we applied morphology, bilateral and median filters to reduce the noise that comes from the hair or any injury on the dorsal hand vein pattern. Morphology filter is useful in closing small holes inside the foreground objects. Bilateral filter removes the noise without smoothing the borders of the veins. Median filter rearranges the values of the pixels in the mask and return the median value. Then, we used contrast limited adaptive histogram equalization (CLAHE), as it is used to apply local histogram equalization on images (see fig.2&3), but with using limited range for histogram to show and keep the main details in the images.

C. Feature Extraction and Selection

We used three features (LBP, LTP and GLCM). LBP depends on the comparison of each pixel in its mask to the center pixel, and if a pixel around it has a bigger value, this means we will add 1 and vice versa. After creating a binary value, we change it to decimal value, which will be the value of the pixel in the new image. LTP has the same steps of LBP, but the surround pixel must be bigger than the center value by certain value.

Table 1. Gender results from the fusion of classifiers

	Cross Validation/Feature vectors	1	2	3	Best K
Gender Fusion KNN-SVM Data set (50%)	Dissimilarity	50	52	56	1-5-11
	Correlation	51	53	55	11-7-3
	Contrast	54	54	56	5-3-9
	ASM	57.5	61.5	63	9-13-9
	LBP	79	83	87	47-37-35
	LTP	63	70	80	5-47-25
Gender Fusion KNN-SVM Data set (50%)	Dissimilarity	50	52	56	1-5-11
	Correlation	51	53	55	11-7-3
	Contrast	54	54	56	5-3-9
	ASM	79.8	72.5	87	11-13-13
	LBP	64.7	70	72	47-37-37
	LTP	67.5	72.6	72	9-47-25

Table 2. Age results from the fusion of classifiers

	Cross Validation/Feature vectors	1	2	3	Best K
Gender Fusion KNN-SVM Data set (50%)	Dissimilarity	50	52	56	1-5-11
	Correlation	51	53	55	11-7-3
	Contrast	54	54	56	5-3-9
	ASM	57.5	61.5	63	9-13-9
	LBP	79	83	87	47-37-35
	LTP	63	70	80	5-47-25
Gender Fusion KNN-SVM Data set (50%)	Dissimilarity	50	52	56	1-5-11
	Correlation	51	53	55	11-7-3
	Contrast	54	54	56	5-3-9
	ASM	79.8	72.5	87	11-13-13
	LBP	64.7	70	72	47-37-37
	LTP	67.5	72.6	72	9-47-25

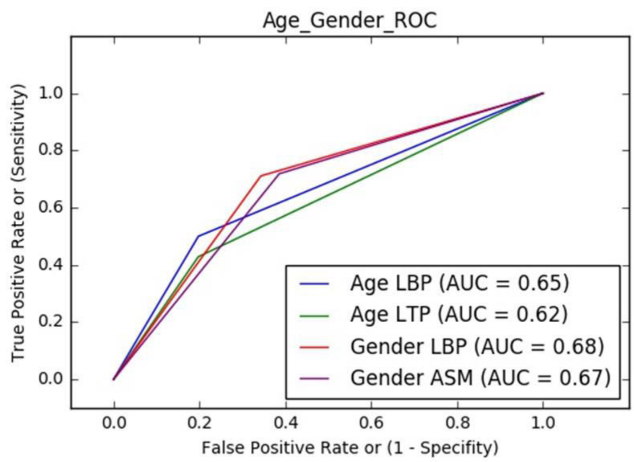

Fig. 4 ROC in age and gender classification after the fusion of classifiers

We mainly focused on the properties of GLCM (contrast, correlation, Angular Second Moment (ASM), and dissimilarity). GLCM is defined as a matrix that is defined over an image to be the distribution of co-occurring pixel values at a given offset.

Then, we used k nearest neighbor (KNN). KNN is a useful technique that assigns weight to the contributions of the neighbors, so that the nearer neighbors contribute more to the average than the more distant ones. KNN is based on Euclidean distance (the distance between new feature vector and surround feature vectors).

We used support vector machine (SVM) as it is based on the idea of finding a hyperplane that best divides a dataset into two classes using support vectors. They are the data points nearest to the hyperplane, the points of a data set that, if removed, would alter the position of the dividing hyperplane. Because of this, they can be considered the critical elements of a data set.

D. Fusion of Classifiers & Features

We used VotingClassifier from Sklearn library in python. This function had the ability to create a proper fusion between KNN and SVM classifiers by giving weights to their decisions based on the probability of each decision. We applied the fusion in the two experiments. We created a fusion between GLCM only, LBP/LTP only or GLCM/LBP/LTP features.

III. Results & Discussion

We applied two experiments that contain 3 K-fold cross validation by using KNN and SVM classifiers. We used LBP, LTP and GLCM features and we normalized the feature vectors.

We got best results from the fusion of KNN and SVM classifiers in age and gender classification. In gender (see table 1), the best result was 85% in 50% data set with best K=35 with LBP and the best result in 70% data set was 85% with best K=13 with ASM. In age (see table 2), the best result in

Table 3. Gender results from the fusion of features

	N times/ Classifiers	1	2	3	Best K
Gender GLCM	KNN	78.5	81.5	82	3-3-3
	SVM	57	63	65	---
Gender LBP/LTP	KNN	74	76	79	49-49-41
	SVM	57	62	63.5	---
Gender GLCM/LBP	KNN	75	77	78	49-37-49
	SVM	56	61	62	---
Gender GLCM/LTP	KNN	65	68	69	3-45-25
	SVM	56	61	65	---
Gender GLCM/LBP/LTP	KNN	74	76	79	49-37-49
	SVM	54	60	63	---

Table 4. Age results from the fusion of features

	N times/ Classifiers	1	2	3	Best K
Age GLCM	KNN	86	85	87	21-21-21
	SVM	80	81	79.8	---
Age LBP/LTP	KNN	79	83	84	43-49-47
	SVM	78	79	78.5	---
Age GLCM/LBP	KNN	78	81	82	41-49-49
	SVM	78.5	79.5	79	---
Age GLCM/LTP	KNN	76	78.5	79	43-47-49
	SVM	79	80	79.5	---
Age GLCM/LBP/LTP	KNN	80	82.5	83	41-49-49
	SVM	79	80	82	---

50% data set was 85% with best K=41 & 47 respectively with LBP and LTP and the best result in 70% data set was 86% with best K=41 & 47 respectively with LBP and LTP. We also got best results from the fusion of GLCM features in age and gender classification. In gender (see table 3), the best result in

KNN with GLCM was 82% and K=3. In Age (see table 4), the best result in KNN with GLCM was 87% with best k=21.

When we used VotingClassifier function with KNN and SVM, we got better results. This function simply gives weight to the decision of each classifier to get a better result. LBP and LTP are the best features that could be extracted from images and ASM is one of the best properties that can be extracted from GLCM, especially after normalization, because this had reduced the error of differences between feature vectors.

In the ROC curve the true positive rate (Sensitivity) is plotted in function of the false positive rate (100-Specificity) for different cut-off points of a parameter. Each point on the ROC curve represents a sensitivity/specificity pair corresponding to a particular decision threshold. The area under the ROC curve (AUC) is a measure of how well a parameter can distinguish between two diagnostic groups (diseased/normal). The ROC of fusion of classifiers has shown that LBP is slightly better than LTP in age and LBP is slightly than ASM in gender due to the differences in the AUC (as shown in Fig.4).

IV. Conclusion

In this paper, we presented a more efficient biometric technique to classify age and genders of individuals using dorsal hand veins. We used a dataset of 1000 images that

represent the left hand of 160 persons (80 males and 80 females, 110 young and 50 old). In age classification, we consider the young people as below 35 years old and the old people as above it. We made two experiments. In the first one we used 70% of the data as a training set as it represents 112 persons or 700 images. This represents 56 males (350 images) and 56 females (350 images) in gender. In the second one we used 50% of the data as a training set as it represents 80 persons (500 images). We used morphology, bilateral and median filter to remove the noise and CLAHE to emphasis the appearance of the dorsal hand veins. We used adaptive threshold to show the veins as black curves. We applied LBP, LTP and GLCM features with KNN and SVM classifiers. We applied VotingClassifier function to create a fusion between features and classifiers. In the fusion of classifiers, we got best results from the fusion of KNN and SVM classifiers in age and gender classification. In gender, the best result was 85% in 50% data set with best K=35 with LBP and the best result in 70% data set was 85% with best K=13 with ASM. In age, the best result in 50% data set was 83% with best K=41 with LBP and LTP and the best result in 70% data set was 84% with best K=47 with LBP and LTP.

IIV. Acknowledgment

We would like to thank our supervisor, Dr. Ahmed M. Badawi, for guiding us in this project. We would like to thank Eng. Menna Essam for helping us in acquiring the data.

Eng. Marina F. Emil, Sarah M. Ragab and Yasmin M. Khattab for collecting the data.

IIIV. Future work

To improve the result of classification among different age ranges, more images need to be acquired from different volunteers with different ages. Add some features to our GUI to make it more user friendly and with extra possibilities. Developing the system with new cameras to get higher resolution images with different positions of the hands. Putting into consideration that BMI would affect and change the pattern of dorsal hand veins, so it would be better to set new parameters to get better results. Acquiring huge data will help in classifying between different ranges of age.

IX. References

[1] Qinghan Xiao, “Security issues in biometric authentication”, 2005.

[2] Mohamed Shahin, Ahmed Badawi, and Mohamed Kamel, "Biometric Authentication Using Fast Correlation of Near Infrared Hand Vein Patterns”, 2008.

[3] Gholamreza Amayeh, George Bebis, Mircea Nicolescu, "Gender Classification from Hand Shape", 2008.

[4] Wafa Damak, Randa Boukhris Trabelsi, Alima Damak Masmoudi, Dorra Sellami, and Amine Nait-Ali, "Age and gender classification from finger vein patterns", 2016.

[5] D. McFadden and E. Shubel.,"Relative lengths of fingers and toes in human males and females", 2002.

[6] Lingyu Wang a, Graham Leedham b, David Siu-Yeung Cho, "Minutiae feature analysis for infra-red hand vein pattern biometric", 2013.

[7] Mihaela Antonina Calin, Sorin Viorel Parasca, Marian Romeo Calin, Emil Petrescu, “An Analysis of Human Dorsal Hand Skin Texture Using Hyperspectral Imaging Technique for Assessing the Skin Aging Process”, 2016.

[8] K Kröger, C Ose, G Rudofsky, J Roesener, D Weiland, H Hirche, “Peripheral veins: influence of gender, body mass index, age and varicose veins on cross-sectional area” ,2003.